

DISCRETIZED MAXIMUM LIKELIHOOD AND ALMOST OPTIMAL
ADAPTIVE CONTROL OF ERGODIC MARKOV MODELS*T. E. DUNCAN[†], B. PASIK-DUNCAN[†], AND L. STETTNER[‡]

Abstract. Three distinct controlled ergodic Markov models are considered here. The models are a discrete time controlled Markov process with complete observations, a controlled diffusion process with complete observations, and a discrete time controlled Markov process with partial observations. The partial observations for the third model have the special form of complete observations in a fixed recurrent set and noisy observations in its complement. For each of the models an almost self-optimizing adaptive control is given. These adaptive controls are constructed from a family of estimates that use a finite discretization of the parameter set and a finite family of almost optimal ergodic controls by a randomized certainty equivalence method. A continuity property of the information of a model for one parameter value with respect to another is used to establish this almost optimality property.

Key words. adaptive control, ergodic control, Markov processes, controlled Markov processes, almost optimal adaptive control

AMS subject classifications. 93E35, 93C40, 60J05, 62M05

PII. S0363012996298369

1. Introduction. In many control problems the models are not completely described and there are perturbations or unmodeled dynamics that are described by noise so that the models are stochastic. If some distributions or parameters in the models are unknown then these control problems can be considered as problems of stochastic adaptive control. In this paper, three unknown ergodic Markov models are considered. The models are a discrete time controlled Markov process with complete observations, a controlled diffusion process with complete observations, and a discrete time controlled Markov process with partial observations. The discrete time Markov processes evolve in a compact state space, and the transition densities depend on an unknown parameter. The partial observations of the discrete time Markov process in the third model have the special form of complete observations in a fixed recurrent set and noisy observations in its complement. The controlled diffusion is described by a stochastic differential equation where the unknown parameter appears in the drift vector. The solution of the stochastic differential equation is given in the weak sense. Since there are some basic differences among these three models, it is convenient to treat them separately. Typically, the results that are given here are stated for each of the three models.

Since the true value of the parameter is unknown, it is estimated using the maximum likelihood procedure where the time differences between the successive updates of the estimates are sufficiently large so that an ergodic property of the information and the cost can be used. Since only almost self-optimality is desired, the maximum likelihood procedure is restricted to choosing from a finite set of possible values for

*Received by the editors February 9, 1996; accepted for publication (in revised form) December 11, 1996. This research was supported by the Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania.

<http://www.siam.org/journals/sicon/36-2/29836.html>

[†]Department of Mathematics, University of Kansas, Lawrence, KS 66045 (duncan@math.ukans.edu, bozenna@kuhub.cc.ukans.edu). This research was supported in part by NSF grants DMS 9305936 and DMS 9623439.

[‡]Institute of Mathematics, Sniadeckich 8, 00–950 Warsaw, Poland (stettner@impan.gov.pl).

the parameter that is a discretization of the possible parameter values. The adaptive strategy uses a randomized certainty equivalence control that chooses with probability almost 1 the control that is almost optimal for the current value of the estimates, and with small, positive probability each of the other almost optimal controls. This procedure is shown to give an almost self-optimizing adaptive control.

The adaptive control of ergodic Markov models has been considered elsewhere (e.g., [1, 2, 4, 6, 8, 9, 10, 13]). However, only here is the maximization of the likelihood function restricted to a finite, discretized set of the possible parameter values. The work of Agrawal [1] has motivated the use here of information and the randomized certainty equivalence adaptive control. A cost-biased maximum likelihood method introduced in [13] is used in [6, 8] for two of the models considered here. The methods used here relax some of the assumptions in [4, 6, 8]. For example, the global Lipschitz continuity of the drift vector with respect to the unknown parameter for the controlled diffusion model is replaced by only continuity, and the requirement that the law of large numbers for some martingales be uniform in the parameter, which necessitated some assumptions in [6, 8], is not required.

The three models that are considered here can be generalized in various ways. The discrete time Markov process can be modified to include the discrete time recursive model in [17]. The controlled diffusion model can be generalized by analogy to [7] to include processes that satisfy stochastic differential equations with delays. The partial observations structure used here can be modified to noisy observations everywhere if there is a sequence of random times such that the process at these times is a family of independent, identically distributed random variables.

The three models are specifically described as follows.

Model I—Discrete time controlled Markov process. A Markov process $(X_n, n \in \mathbb{N})$ evolves in a compact metric space E with the transition operator $P(x_n, dy; v_n, \alpha^0)$ at time $n \in \mathbb{N}$, where $\alpha^0 \in \mathcal{A}$ is an unknown fixed parameter and \mathcal{A} is a compact metric space, and the control v_n takes values in a compact metric space U and is adapted to $\sigma(X_0, \dots, X_n)$. A generic parameter value $\alpha \in \mathcal{A}$ has a transition operator that is described by replacing α^0 by α above. The transition operators have continuous densities with respect to a fixed measure $\varphi(\cdot)$; that is, for each $B \in \mathcal{B}(E)$, the Borel σ -algebra on E , and each $\alpha \in \mathcal{A}$,

$$(1) \quad P(x, B; v, \alpha) = \int_B p(x, y, v, \alpha) \varphi(dy),$$

where φ is a probability measure on E and $p : E \times E \times U \times \mathcal{A} \rightarrow \mathbb{R}_+$ is continuous. It is assumed that $p(x, y, v, \alpha) > 0$ for all $x, y \in E$, $v \in U$, and $\alpha \in \mathcal{A}$, and $\text{supp } \varphi = E$. The control problem is to minimize the following ergodic cost functional:

$$(2) \quad I^1((v_n, n \in \mathbb{N})) = \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} c(X_i, v_i),$$

where $c : E \times U \rightarrow \mathbb{R}_+$ is a bounded, Borel measurable function. The family of controls $(v_n, n \in \mathbb{N})$ has the form $v_n = u(X_n)$, where $u \in \mathcal{U}$ and \mathcal{U} is the family of Borel measurable functions from E to U .

Model II—Controlled diffusion process. Let $(X(t), t \in \mathbb{R}_+)$ be a controlled diffusion process that satisfies the following stochastic differential equation:

$$(3) \quad \begin{aligned} dX(t) &= f(X(t))dt + h(X(t), \alpha^0, v(t))dt + \sigma(X(t))dW(t), \\ X(0) &= x, \end{aligned}$$

where $X(t) \in \mathbb{R}^n$, $(W(t), t \geq 0)$ is a standard \mathbb{R}^n -valued Wiener process, $\alpha^0 \in \mathcal{A}$ is unknown and \mathcal{A} is compact, $(v(t), t \geq 0)$ is adapted to $\sigma(X(s), s \leq t)$, and $v(t) \in U$, a compact set. The functions f and σ satisfy a global Lipschitz condition, $\sigma(x)\sigma^*(x) \geq cI > 0$ for all $x \in \mathbb{R}^n$, and $h : \mathbb{R}^n \times \mathcal{A} \times U \rightarrow \mathbb{R}^n$ is a bounded Borel measurable function and either $h(x, \cdot, v)$ is continuous uniformly in $v \in U$ and x from compact subsets of \mathbb{R}^n or $h(x, \cdot, \cdot)$ is continuous for each $x \in \mathbb{R}^n$. The solution of the stochastic differential equation (3) is given in the weak sense by an absolutely continuous transformation of the measure of the strong solution of

$$(4) \quad \begin{aligned} dY(t) &= f(Y(t))dt + \sigma(Y(t))dW(t), \\ Y(0) &= x. \end{aligned}$$

The family of controls $(v(t), t \geq 0)$ has the form $v(t) = u(X(t))$, where $u \in \mathcal{U}$ and \mathcal{U} is the family of Borel measurable functions from \mathbb{R}^n into U . Let T_A be the first hitting time of $A \in \mathcal{B}(\mathbb{R}^n)$; that is,

$$T_A = \begin{cases} \inf\{s > 0 : X(s) \in A\}, \\ +\infty \end{cases} \quad \text{if the above set is empty.}$$

Let Γ_1 and Γ_2 be two spheres in \mathbb{R}^n with centers at 0 and radii $0 < r_1 < r_2$, respectively. Let τ be given as

$$(5) \quad \tau = T_{\Gamma_2} + T_{\Gamma_1} \circ \theta_{T_{\Gamma_2}},$$

where $(\theta_t, t \geq 0)$ is the family of shift operators acting on $C(\mathbb{R}_+, \mathbb{R}^n)$. The random time τ is the first time that the process $(X(t), t \geq 0)$ hits Γ_1 after hitting Γ_2 . It is assumed that

$$(6) \quad \sup_{\alpha \in \mathcal{A}} \sup_{u \in \mathcal{U}} \sup_{x \in \Gamma_1} E_x^{\alpha, u}[\tau^2] < \infty$$

and

$$(7) \quad E_x^{\alpha, u}[T_{\Gamma_1}] < \infty$$

for each $(x, \alpha, u) \in \mathbb{R}^n \times \mathcal{A} \times \mathcal{U}$, where $E_x^{\alpha, u}$ is the expectation with respect to a process $(X(t), t \geq 0)$ that satisfies (3) with α^0 replaced by α and $v(t) = u(X(t))$. The dependence of the solution of (3) on $\alpha \in \mathcal{A}$ and the control $(v(t), t \geq 0)$ is suppressed for notational convenience. However, it is shown explicitly when the expectations of functions of the solution are taken. The control problem is to minimize the ergodic cost functional

$$(8) \quad I^2((v(t), t \geq 0)) = \limsup_{t \rightarrow \infty} \frac{1}{t} \int_0^t c(X(s), v(s))ds,$$

where $c : \mathbb{R}^n \times U \rightarrow \mathbb{R}_+$ is a bounded, Borel measurable function.

Model III—A partially observed discrete time controlled Markov process. A controlled Markov process $(X_n, n \in \mathbb{N})$ evolves in a compact subset E of \mathbb{R}^d with the transition operator $P(x_n, dy; v_n, \alpha^0)$ at time $n \in \mathbb{N}$, where $\alpha^0 \in \mathcal{A}$ is an unknown parameter, \mathcal{A} is a compact metric space, and the control v_n takes values in a compact metric space U . A generic parameter value $\alpha \in \mathcal{A}$ has a transition operator

that is described by replacing α^0 by α above. The transition operators have densities with respect to Lebesgue measure; that is,

$$(9) \quad P(x, B; v, \alpha) = \int_B p(x, y, v, \alpha) dy,$$

where $\alpha \in \mathcal{A}$, $B \in \mathcal{B}(E)$, and $p : E \times E \times U \times \mathcal{A} \rightarrow \mathbb{R}_+$ is continuous and $p > 0$ on $E \times E \times U \times \mathcal{A}$. The process $(X_n, n \in \mathbb{N})$ is completely observed in a nonempty compact subset $\Gamma \subset E$ and is partially observed in $E \setminus \Gamma$. The observation process $(Y_n, n \in \mathbb{N})$ is explicitly described as follows:

$$(10) \quad P(Y_i \in B | X_i, \mathcal{Y}_{i-1}) = 1_{B \cap \Gamma}(X_i) + 1_{\Gamma^c}(X_i) \int_{B \cap \Gamma^c} r(X_i, y) dy,$$

where $B \in \mathcal{B}(E)$, $\mathcal{Y}_i = \sigma(Y_1, \dots, Y_i)$, $\mathcal{Y}_0 = \{\emptyset, \Omega\}$, and $r : \Gamma^c \times \Gamma^c \rightarrow \mathbb{R}$ is Borel measurable such that $\int_{\Gamma^c} r(x, y) dy = 1$ for each $x \in \Gamma^c$. A control v is a U -valued, \mathcal{Y}_n -adapted process. The control problem is to minimize the ergodic cost functional

$$(11) \quad I^3((v_n, n \in \mathbb{N})) = \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} c(X_i, v_i),$$

where $c \in C(E \times U)$. It is assumed that there is a nonempty compact set $\Gamma_1 \subset \Gamma$ such that for each probability law μ on X_0 , control $u = (v_n, n \in \mathbb{N})$, and $\alpha \in \mathcal{A}$,

$$(12) \quad E_\mu^{\alpha, u}[T_{\Gamma_1}] < \infty$$

and

$$(13) \quad \sup_{x \in \Gamma_1} \sup_{u \in \mathcal{U}} E_x^{\alpha, u}[\tau^2] < \infty,$$

where T_{Γ_1} is the first hitting time of Γ_1 , τ is the first hitting time of Γ_1 after hitting Γ^c ($\tau = T_{\Gamma^c} + T_{\Gamma_1} \circ \theta_{T_{\Gamma^c}}$), and $E_\mu^{\alpha, u}$ is the expectation for the process $(X_n, n \in \mathbb{N})$ with initial law μ , control u , and parameter $\alpha \in \mathcal{A}$. For a probability law μ for X_0 the measure-valued process $(\Pi_n^{\alpha^0}, n \in \mathbb{N})$ is defined as follows:

$$(14) \quad \Pi_0^{\alpha^0}(B) = \mu(B),$$

$$(15) \quad \Pi_n^{\alpha^0}(B) = P_n(X_n \in B | \mathcal{Y}_n)$$

for each $B \in \mathcal{B}(E)$. This conditional measure process can be represented more explicitly using (10) (e.g., Lemma 1 of [18]) as follows:

$$(16) \quad \Pi_{n+1}^{\alpha^0}(B) = 1_{B \cap \Gamma}(Y_{n+1}) + 1_{\Gamma^c}(Y_{n+1}) M(Y_{n+1}, \Pi_n^{\alpha^0}, v_n, \alpha^0)(B),$$

where

$$(17) \quad M(y, \nu, v, \alpha^0) = \frac{\int_{B \cap \Gamma^c} r(z, y) p(\nu, z, v, \alpha^0) dz}{\int_{\Gamma^c} r(z, y) p(\nu, z, v, \alpha^0) dz}$$

and

$$(18) \quad p(\nu, z, v, \alpha^0) = \int_E p(x, z, v, \alpha^0) \nu(dx).$$

2. A finite family of almost optimal controls. For the adaptive control of Models I, II, and III a finite family of controls is constructed that includes at least one that is almost optimal for each parameter value $\alpha \in \mathcal{A}$.

To determine the almost optimal controls the averaged versions of the ergodic cost functionals (2), (8), (11) are used. These are denoted as follows:

$$(19) \quad J_{\mu}^{\alpha^0,1}((v_n, n \in \mathbb{N})) = \limsup_{n \rightarrow \infty} \frac{1}{n} E_{\mu}^{\alpha^0,v} \left[\sum_{i=0}^{n-1} c(X_i, v_i) \right],$$

$$(20) \quad J_{\mu}^{\alpha^0,2}((v(t), t \geq 0)) = \limsup_{t \rightarrow \infty} \frac{1}{t} E_{\mu}^{\alpha^0,v} \left[\int_0^t c(X(s), v(s)) ds \right],$$

$$(21) \quad J_{\mu}^{\alpha^0,3}((v_n, n \in \mathbb{N})) = \limsup_{n \rightarrow \infty} \frac{1}{n} E_{\mu}^{\alpha^0,v} \left[\sum_{i=0}^{n-1} c(X_i, v_i) \right],$$

where μ is the probability law for X_0 and $\alpha^0 \in \mathcal{A}$ is the true parameter value. The finite families of almost optimal controls for Models I, II, and III are constructed for the cost functionals $J_{\mu}^{\alpha^0,1}$, $J_{\mu}^{\alpha^0,2}$, and $J_{\mu}^{\alpha^0,3}$, respectively.

Model I. It is assumed that it suffices to consider controls of the form $v_n = u_n(X_n)$, where $u_n \in \mathcal{U} = B(E, U)$, the family of Borel measurable functions from E to U . Clearly, this restriction is satisfied if c is a continuous, bounded function because by (1) for $B \in \mathcal{B}(E)$, $x \in E$, $v \in U$, and $\alpha \in \mathcal{A}$,

$$(22) \quad P^{v,\alpha}(x, B) \geq \inf_{x,y \in E} \inf_{v \in U} \inf_{\alpha \in \mathcal{A}} p(x, y, v, \alpha) \varphi(B)$$

and

$$\inf_{x,y \in E} \inf_{v \in U} \inf_{\alpha \in \mathcal{A}} p(x, y, v, \alpha) > 0,$$

and (see Theorem 2.2 and Corollary 3.6 in Chap. 3 of [10]) for each $\alpha \in \mathcal{A}$, there is a $u_{\alpha} \in B(E, U)$ such that $J_{\mu}^{\alpha,1}(u_{\alpha})$ is optimal.

For Model I there is a uniform ergodicity property and a finite family of almost optimal controls.

PROPOSITION 1. *For Model I with $\alpha \in \mathcal{A}$ and $u \in \mathcal{U}$ there is a probability measure π_u^{α} on $\mathcal{B}(E)$ such that*

$$(23) \quad \sup_{u \in \mathcal{U}} \sup_{\alpha \in \mathcal{A}} \sup_{x \in E} \|(P^{u,\alpha})^n(x, \cdot) - \pi_u^{\alpha}(\cdot)\|_{\text{var}} \leq 2\gamma_0^{n-1}$$

where $\|\cdot\|_{\text{var}}$ is the variation norm and

$$\gamma_0 = 1 - \inf_{x,y \in E} \inf_{v \in U} \inf_{\alpha \in \mathcal{A}} p(x, y, v, \alpha).$$

There is a constant K_1 such that for $\alpha, \beta \in \mathcal{A}$ and $u \in \mathcal{U}$

$$(24) \quad \|\pi_u^{\alpha} - \pi_u^{\beta}\|_{\text{var}} \leq K_1 \sup_{x \in E} \|P^{u,\alpha}(x, \cdot) - P^{u,\beta}(x, \cdot)\|_{\text{var}}.$$

Furthermore, given $\varepsilon > 0$ there is a finite family of controls $\mathcal{U}^1(\varepsilon) = \{u^1, \dots, u^{r(\varepsilon)}\}$ such that for each $\alpha \in \mathcal{A}$ there is a $k \in \{1, \dots, r(\varepsilon)\}$ and

$$(25) \quad J_{\mu}^{\alpha,1}(u^k(X(\cdot))) \leq \lambda_1(\alpha) + \varepsilon,$$

where

$$(26) \quad \lambda_1(\alpha) = \inf_{u \in \mathcal{U}} J_{\mu}^{\alpha,1}(u(X(\cdot))).$$

Proof. From (22) and (5.6) of [5], (23) is verified. The inequality (24) follows from the proof of Proposition 1 of [17]. The existence of a finite family $\mathcal{U}^1(\varepsilon)$ satisfying (25) follows from (24) and the proof of Lemma 2 of [17]. \square

Model II. Let $(\tau_n, n \in \mathbb{N})$ be an increasing sequence of random times such that $\tau_1 = \tau$ and $\tau_{n+1} = \tau_n \circ \theta_{\tau_n}$ for $n > 1$. For a given control $u \in \mathcal{U}$ and parameter $\alpha^0 \in \mathcal{A}$ there is a unique invariant measure $\eta_u^{\alpha^0}$ for the embedded Markov chain $(X_{\tau_n}, n \in \mathbb{N})$ and $X_0 \in \Gamma_1$ (e.g., [6]). Furthermore, there is a unique invariant measure $\pi_u^{\alpha^0}$ for the process $(X(t), t \geq 0)$ with $v(t) = u(X(t))$, and it has the form

$$(27) \quad \pi_u^{\alpha^0}(B) = \frac{\int_{\Gamma_1} E_x^{\alpha^0, u} \left[\int_0^{\tau} 1_B(X(s)) ds \right] \eta_u^{\alpha^0}(dx)}{\int_{\Gamma_1} E_x^{\alpha^0, u} [\tau] \eta_u^{\alpha^0}(dx)}.$$

For Model II there is an analogue of Proposition 1.

PROPOSITION 2. *For Model II there is $\gamma_0 \in (0, 1)$ such that*

$$(28) \quad \sup_{u \in \mathcal{U}} \sup_{\alpha \in \mathcal{A}} \sup_{x \in \Gamma_1} \sup_{B \in \mathcal{B}(\Gamma_1)} |P_x^{\alpha, u}(X(\tau_n) \in B) - \eta_u^{\alpha}(B)| \leq \gamma_0^n,$$

where η_u^{α} is the unique invariant measure for the embedded Markov chain. There is a constant K_1 such that for $\alpha, \beta \in \mathcal{A}$ and $u \in \mathcal{U}$,

$$(29) \quad \|\eta_u^{\alpha} - \eta_u^{\beta}\|_{\text{var}} \leq K_1 \sup_{x \in E} \sup_{B \in \mathcal{B}(\Gamma_1)} |P_x^{\alpha, u}(X(\tau) \in B) - P_x^{\beta, u}(X(\tau) \in B)|.$$

Furthermore, given $\varepsilon > 0$ there is a $\delta > 0$ such that if $\alpha, \beta \in \mathcal{A}$ and $\rho_{\mathcal{A}}(\alpha, \beta) < \delta$ then

$$(30) \quad \sup_{u \in \mathcal{U}} \sup_{x \in \Gamma_1} \sup_{B \in \mathcal{B}(\Gamma_1)} |P_x^{\alpha, u}(X(\tau) \in B) - P_x^{\beta, u}(X(\tau) \in B)| < \varepsilon$$

and

$$(31) \quad \sup_{u \in \mathcal{U}} \sup_{x \in \Gamma_1} \sup_{B \in \mathcal{B}(\Gamma_1)} \left| E_x^{\alpha, u} \left[\int_0^{\tau} 1_B(X(s)) ds \right] - E_x^{\beta, u} \left[\int_0^{\tau} 1_B(X(s)) ds \right] \right| < \varepsilon,$$

where $\rho_{\mathcal{A}}$ is a metric on \mathcal{A} compatible with its topology. Furthermore, given $\varepsilon > 0$, there is a finite family of controls $\mathcal{U}^2(\varepsilon) = \{u^1, \dots, u^{r(\varepsilon)}\}$ such that for each $\alpha \in \mathcal{A}$ there is a $k \in \{1, \dots, r(\varepsilon)\}$ and

$$(32) \quad J_{\mu}^{\alpha, 2}(u^k(X(\cdot))) \leq \lambda_2(\alpha) + \varepsilon,$$

where

$$(33) \quad \lambda_2(\alpha) = \inf_{u \in \mathcal{U}} J_{\mu}^{\alpha, 2}(u(X(\cdot))).$$

Proof. By Proposition 2.2 of [6] and Theorem 4.1 of [3] the inequality (28) follows. Using the proof of Proposition 1 of [6], as in our Proposition 1, the inequality (29)

follows. The uniform continuity properties (30), (31) can be verified as for (10) and (19) of [6]. Since h is not assumed to be Lipschitz continuous with respect to $\alpha \in \mathcal{A}$, it is necessary to verify that the map

$$(34) \quad H : \Gamma_1 \times \mathcal{A} \rightarrow \mathbb{R}$$

given by

$$(35) \quad H(x, \alpha) = E_x \left[\int_0^t |\sigma^{-1}(Y(s))h(Y(s), \alpha, u(Y(s)))|^2 ds \right]$$

is continuous uniformly in $u \in \mathcal{U}$, where E_x is the expectation for P_x that is the measure for the solution of (4), and

$$E_x \left[\int_0^t |\sigma^{-1}(Y(s))(h(Y(s), \alpha, u(Y(s))) - h(Y(s), \beta, u(Y(s))))|^2 ds \right] \rightarrow 0$$

as $\rho_{\mathcal{A}}(\alpha, \beta) \rightarrow 0$ uniformly in $u \in U$. The proof of this last continuity is similar to the verification of the continuity of H , so only the verification of \bar{H} is given. Since h is bounded, it is sufficient to verify the continuity of $\bar{H} : \Gamma_1 \times \mathcal{A} \rightarrow \mathbb{R}$ given by

$$(36) \quad \bar{H}(x, \alpha) = \int_{t_1}^t E_x |\sigma^{-1}(Y(s))h(Y(s), \alpha, u(Y(s)))|^2 ds$$

for each $t_1 < t$ uniformly in $u \in U$. To verify this continuity note that the map $(s, x) \in (0, \infty) \times \mathbb{R}^d \mapsto P_x(Y(s) \in \cdot)$ is continuous in the variation norm topology. In fact, by Lemma 9.22 of [19], for $s_n \rightarrow s > 0$ and $x_n \rightarrow x$ the family of measures $(P_{x_n}(Y(s_n) \in \cdot), n \in \mathbb{N})$ is tight, so for any $\varepsilon > 0$ there is a compact set $K \subset \mathbb{R}^d$ such that

$$P_y(Y(s_n) \in K^c) < \varepsilon$$

for all $y \in \{x, x_1, x_2, \dots\}$. By Theorem 3.2.1 of [19] the measures $(P(Y(s) \in \cdot), s > 0)$ have continuous densities. As $n \rightarrow \infty$ the following inequality is easily verified from the previous inequality:

$$\sup_{B \in \mathcal{B}(\mathbb{R}^d)} |P_{x_n}(Y(s_n) \in B) - P_x(Y(s) \in B)| \leq 2\varepsilon + \int_K |p(s_n, x_n, y) - p(s, x, y)| dy.$$

Thus the continuity in the variation norm of $P(Y(s) \in \cdot)$ is verified. By this continuity and the continuity of $\alpha \mapsto h(y, \alpha, v)$ that is uniform in $v \in U$, the continuity of (36) and therefore (34) follows. Now only the verification of (32) remains. By (27), (29), (30), (31), given $\varepsilon > 0$, there is a $\delta > 0$ such that if $\alpha, \beta \in \mathcal{A}$ and $\rho_{\mathcal{A}}(\alpha, \beta) < \delta$, then

$$\sup_{u \in \mathcal{U}} \|\pi_u^\alpha - \pi_u^\beta\|_{\text{var}} < \varepsilon.$$

So by Propositions 2.3 and 2.4 of [6] there is a finite family of controls $\mathcal{U}^2(\varepsilon)$ such that the inequality (32) is satisfied. \square

Model III. Let \mathcal{U} be a fixed compact subset of $C(\mathcal{P}(E), U)$ where $P(E)$ is the family of probability measures on E with the vague topology, and let $u(\alpha)$ be the control sequence such that $v_n = u(\Pi_n^\alpha)$. Define λ_3 as follows:

$$(37) \quad \lambda_3(\alpha) = \inf_{u \in \mathcal{U}} J_\mu^{\alpha, 3}(u(\Pi_n^\alpha))$$

and an increasing sequence of random times $(\tau_n, n \in \mathbb{N})$ as $\tau_1 = \tau$, $\tau_{n+1} = \tau_n + \tau \circ \theta_{\tau_n}$, where $(\theta_t, t \geq 0)$ is the family of shift operators acting on $C(\mathbb{R}_+, \mathbb{R}^d)$.

Some additional assumptions are made on Model III.

(A1) The function r given in (10) is continuous on $\Gamma^c \times \Gamma^c$ and bounded on $\Gamma_\delta^c \times \Gamma_\delta^c$ for some $\delta > 0$ where $\Gamma_\delta^c = \{(y, z) \in \Gamma^c \times \Gamma^c : \rho_E(y, \Gamma) \geq \delta\} \cup \{(y, z) \in \Gamma^c \times \Gamma^c : \rho_E(y, z) \geq \delta\}$ and ρ_E is a metric on E that is compatible with its topology. If $(y_n, n \in \mathbb{N})$ is a sequence in Γ^c such that $y_n \rightarrow y \in \Gamma$ as $n \rightarrow \infty$, and for $\delta > 0$, $B(y, \delta) = \{z \in \Gamma^c : \rho_E(z, y) \leq \delta\}$, then

$$(38) \quad \lim_{n \rightarrow \infty} \inf_{\alpha \in A} \inf_{v \in \mathcal{U}} \inf_{x \in K} \int_{B(y, \delta)} r(z, y_n) P^{\alpha, v}(x, dz) = \infty$$

for any compact subset $K \subset E$.

(A2) If $(z_n, n \in \mathbb{N})$ is a sequence in Γ^c that converges to z , then

$$\lim_{n \rightarrow \infty} R(z_n, \cdot) = R(z, \cdot),$$

where the topology is the vague convergence of measures and

$$(39) \quad R(z, A) = \begin{cases} \int_{A \cap \Gamma^c} r(z, y) dy & \text{for } z \in \Gamma^c, \\ 1_A(z) & \text{for } z \in \Gamma \end{cases}$$

for $A \in \mathcal{B}(E)$.

Using (A1) and (A2) an analogue of Propositions 1 and 2 is given for Model III.

PROPOSITION 3. *For Model III, if (A1) and (A2) are satisfied, then there is a $\gamma_0 \in (0, 1)$ and a measure $\eta_{u(\beta)}^\alpha$ on Γ such that*

$$(40) \quad \sup_{u \in \mathcal{U}} \sup_{\alpha, \beta \in A} \sup_{x \in \Gamma} \sup_{B \in \mathcal{B}(\Gamma_1)} |P_x^{\alpha, u(\beta)}(X_{\tau_n} \in B) - \eta_{u(\beta)}^\alpha(B)| < \gamma_0^n.$$

Given $\varepsilon > 0$ there is a finite family of controls $\tilde{\mathcal{U}}(\varepsilon) = \{u^1, \dots, u^{r(\varepsilon)}\} \subset \tilde{\mathcal{U}}$ and $\delta_0 > 0$ such that if $\rho_A(\alpha, \beta) < \delta_0$, then

$$(41) \quad \lambda_3(\beta) - \varepsilon \leq J_\mu^{\beta, 3}(u^k(\Pi_n^\alpha)) \leq \lambda_3(\beta) + \varepsilon$$

for some $k \in \{1, \dots, r(\varepsilon)\}$.

Proof. The continuity and the positivity of the transition density p and (13) imply (40). Using the proof of Lemma 3 of [17], it follows that

$$J_\mu^{\beta, 3}(u^k(\Pi_n^\alpha)) = \frac{\int_\Gamma E_x^{\beta, u(\alpha)} \left[\sum_{i=0}^{\tau-1} \int_E c(z, u(\Pi_i^\alpha)) \Pi_i^\beta(dz) \eta_{u(\alpha)}^\beta(dx) \right]}{\int_\Gamma E_x^{\beta, u(\alpha)} [\tau] \eta_{u(\alpha)}^\beta(dx)}.$$

By the proofs of Lemma 2 of [17] and Proposition 2.4 of [6], for the verification of (41) it suffices to show that given $\varepsilon > 0$, there is a $\delta > 0$ such that for $\alpha, \beta, \alpha_1, \beta_1 \in A$, if $\rho_A(\alpha, \alpha_1) < \delta$ and $\rho_A(\beta, \beta_1) < \delta$ then

$$(42) \quad \sup_{u \in \mathcal{U}} |J_\mu^{\beta, 3}(u(\Pi_n^\alpha)) - J_\mu^{\beta_1, 3}(u(\Pi_n^{\alpha_1}))| < \varepsilon.$$

If the inequality (42) is not satisfied then there are sequences $(\alpha^m, m \in \mathbb{N})$, $(\alpha_1^m, m \in \mathbb{N})$ such that $\alpha^m \rightarrow \alpha$, $\alpha_1^m \rightarrow \alpha_1$, and $u_m(v) \rightarrow u(v)$ uniformly in $v \in \mathcal{P}(E)$ as $m \rightarrow \infty$ and

$$(43) \quad |J_\mu^{\alpha^m, 3}(u_m(\Pi_n^{\alpha^m})) - J_\mu^{\alpha_1^m, 3}(u_m(\Pi_n^{\alpha_1^m}))| \geq \varepsilon > 0$$

for all $m \in \mathbb{N}$. Using some continuity arguments in the proofs of Theorems 1 and 6 of [18], where the pair $(\alpha, v) \in \mathcal{A} \times U$ is considered as the control, it follows that

$$\lim_{m \rightarrow \infty} J_\mu^{\alpha^m, 3}(u_m(\Pi_n^{\alpha^m})) = J_\mu^{\alpha, 3}(u(\Pi_n^\alpha))$$

and

$$\lim_{m \rightarrow \infty} J_\mu^{\alpha_1^m, 3}(u_m(\Pi_n^{\alpha_1^m})) = J_\mu^{\alpha, 3}(u(\Pi_n^\alpha)),$$

which contradict the inequality (43). Thus (42) is satisfied and there is a finite family $\tilde{\mathcal{U}}(\varepsilon)$ of controls such that (41) is satisfied. \square

Let $\tilde{\mathcal{A}}(\delta_0) = \{\alpha(1), \dots, \alpha(k(\delta_0))\}$ be distinguished points, one from each of a finite δ_0 net in \mathcal{A} . By Proposition 3, given $\varepsilon > 0$, there is an $\tilde{\mathcal{A}}(\delta_0)$ from a δ_0 net of \mathcal{A} such that the controls $(u^k(\Pi_n^\alpha), k \in \{1, \dots, r(\varepsilon)\})$ and $\alpha \in \tilde{\mathcal{A}}(\delta_0)$ form the family of ε optimal controls.

Remark. A finite family of controls for $\mathcal{U}^{(1)}(\varepsilon)$ can be obtained from a discretization of the Bellman equation (cf. [6] and Section 3.5 of [10]). A finite family of controls for $\mathcal{U}^{(2)}(\varepsilon)$ can be obtained using [14], and a finite family of controls for $\tilde{\mathcal{U}}(\varepsilon)$ can be obtained using [15].

3. The information for different parameters. Kullback and Leibler [11] have used a notion of information in statistics. For the adaptive control problems for the three models considered here the information is computed from the probability densities for different values of the unknown parameter. It is described in [16] as the information of one parameter value with respect to another. It is shown in [12] that it is naturally related to the notion of information in information theory. This quantity has a different form for each of the Markov models. It is denoted K^i , $i = 1, 2, 3$, for the three models.

$$(44) \quad K_u^1(\alpha, \beta) = \int_E \int_E \ln \left(\frac{p(x, y, u(x), \beta)}{p(x, y, u(x), \alpha)} \right) p(x, y, u(x), \beta) \varphi(dy) \pi_u^\beta(dx),$$

where φ is given in (1), π is the invariant measure given in (23), $\alpha, \beta \in \mathcal{A}$, and $u \in \mathcal{U}$.

$$(45) \quad \begin{aligned} K_u^2(\alpha, \beta) &= \frac{1}{2} \int_{\Gamma_1} E_x^{\beta, u} \left[\int_0^\tau |\sigma^{-1}(X(s))(h(X(s), \alpha, u(X(s))) - h(X(s), \beta, u(X(s))))|^2 ds \right] \\ &\quad \cdot \eta_u^\beta(dx) \left(\int_{\Gamma_1} E_x^{\beta, u}[\tau] \eta_u^\beta(dx) \right)^{-1}, \end{aligned}$$

where η is the invariant measure for the embedded Markov chain given in (27), $\alpha, \beta \in \mathcal{A}$, and $u \in \mathcal{U}$.

$$K_u^3(\alpha, \beta, \gamma) = \int_\Gamma E_x^{\beta, u(\gamma)} \left[\sum_{i=0}^{\tau-1} \ln \left(\frac{F(\Pi_i^\beta, u(\Pi_i^\gamma), \beta)(Y_{i+1})}{F(\Pi_i^\alpha, u(\Pi_i^\gamma), \alpha)(Y_{i+1})} \right) \right]$$

$$(46) \quad \cdot \eta_{u(\gamma)}^\beta(dx) \left(\int_{\Gamma} E_x^{\beta, u(\gamma)}[\tau] \eta_{u(\gamma)}^\beta(dx) \right)^{-1},$$

where η is given in (40), $\alpha, \beta, \gamma \in \mathcal{A}$, $u \in \tilde{\mathcal{U}}$, $u(\gamma)$ in (46) indicates that the control $u(\Pi_i^\gamma)$ is used, and

$$(47) \quad \begin{aligned} F(\nu, v, \alpha)(y) &:= 1_{\Gamma}(y) p(\nu, y, v, \alpha) \\ &+ 1_{E \setminus \Gamma}(y) \int_{E \setminus \Gamma} r(z, y) p(\nu, z, v, \alpha) dz. \end{aligned}$$

Now some important properties are verified for K^i , $i = 1, 2, 3$.

PROPOSITION 4. *Consider Model I with the assumptions imposed on it. For each $u \in \mathcal{U}$ the map $K_u^1 : \mathcal{A} \times \mathcal{A} \rightarrow \mathbb{R}$ is continuous. Furthermore, if $K_u^1(\alpha, \beta) = 0$, then $\pi_u^\alpha = \pi_u^\beta$.*

Proof. Let $L : \mathcal{A} \times \mathcal{A} \times E \rightarrow \mathbb{R}$ be given by

$$L(\alpha, \beta, x) = \int_E \ln \left(\frac{p(x, y, u(x), \beta)}{p(x, y, u(x), \alpha)} \right) p(x, y, u(x), \beta) \varphi(dy).$$

$L(\cdot, \cdot, x)$ is continuous and bounded uniformly in $x \in E$, so the continuity of K_u^1 follows by (24).

For $x \in E$, Jensen's inequality implies that

$$\int_E \ln \left(\frac{p(x, y, u(x), \beta)}{p(x, y, u(x), \alpha)} \right) p(x, y, u(x), \beta) \varphi(dy) \geq 0.$$

For each $B \in \mathcal{B}(E)$ it follows by the definition of invariant measures that

$$\pi_u^\beta(B) \geq \inf_{x, y \in E} \inf_{v \in U, \beta \in \mathcal{A}} p(x, y, v, \beta) \varphi(B).$$

If $K_u^1(\alpha, \beta) = 0$, then

$$(48) \quad \int_E \ln \left(\frac{p(x, y, u(x), \beta)}{p(x, y, u(x), \alpha)} \right) p(x, y, u(x), \beta) \varphi(dy) = 0$$

for (φ) almost all $x \in E$. Since $\ln(\cdot)$ is a strongly convex function it follows by Jensen's inequality that

$$(49) \quad p(x, y, u(x), \alpha) = p(x, y, u(x), \beta)$$

for (φ) almost all $x \in E$ and (φ) almost all $y \in E$. Thus for $B \in \mathcal{B}(E)$,

$$\begin{aligned} \pi_u^\beta(B) &= \int_E \int_B p(x, y, u(x), \beta) \varphi(dy) \pi_u^\beta(dx) \\ &= \int_E \int_B p(x, y, u(x), \alpha) \varphi(dy) \pi_u^\beta(dx) \\ &= \int_E P^{u, \alpha}(x, B) \pi_u^\beta(dx). \end{aligned}$$

The last equality implies that π_u^β is an invariant measure for the transition operator $P^{\alpha, u}$. The uniqueness of the invariant measure for $P^{\alpha, u}$, which follows from (23), implies that $\pi_u^\beta = \pi_u^\alpha$. \square

Now a result analogous to the above proposition is obtained for Model II.

PROPOSITION 5. Consider Model II with the assumptions imposed on it. For each $u \in \mathcal{U}$, the map $K_u^2 : \mathcal{A} \times \mathcal{A} \rightarrow \mathbb{R}$ is continuous. Furthermore, if $K_u^2(\alpha, \beta) = 0$, then $\pi_u^\alpha = \pi_u^\beta$.

Proof. The continuity of K_u^2 follows from (29), (30), (31) and the continuity of $h(x, \cdot, v)$. If $K_u^2(\alpha, \beta) = 0$, then by Lemma 3.4 of [6] it follows that

$$h(x, \alpha, u(x)) = h(x, \beta, u(x))$$

for all $x \in \mathbb{R}^n \setminus D$ where $\lambda(D) = 0$ and λ is an n -dimensional Lebesgue measure. Thus for (λ) almost all $x \in \mathbb{R}^n$, $t > 0$, and $B \in \mathcal{B}(\mathbb{R}^n)$,

$$P_x^{\alpha, u}(X(t) \in B) = P_x^{\beta, u}(X(t) \in B).$$

The uniqueness of the invariant measures, as in the proof of Proposition 4, implies that $\pi_u^\alpha = \pi_u^\beta$. \square

For Model III an additional assumption is introduced:

$$(A5) \quad \int_{E \setminus \Gamma} r(x, y)(p(\nu_1, x, v, \alpha) - p(\nu_2, x, v, \beta)) dx = 0$$

for almost all $y \in E \setminus \Gamma$ if and only if

$$p(\nu_1, x, v, \alpha) = p(\nu_2, x, v, \beta)$$

for almost all $x \in E \setminus \Gamma$.

Now an analogue of the previous two propositions is verified for Model III.

PROPOSITION 6. If (A1)–(A5) are satisfied for Model III, then for each $u \in \tilde{\mathcal{U}}$ the map $K_u^3 : \mathcal{A} \times \mathcal{A} \times \mathcal{A} \rightarrow \mathbb{R}$ is continuous. Furthermore, if $K_u^3(\alpha, \beta, \gamma) = 0$ then the measures $\Psi_{u(\gamma)}^\alpha$ and $\Psi_{u(\gamma)}^\beta$ on the Borel σ -algebra of $\mathcal{P}(E) \times \mathcal{P}(E)$ coincide, where

$$(50) \quad \Psi_{u(\gamma)}^\delta(B) = \frac{\int_{\Gamma} E_x^{\delta, u(\gamma)} \left[\sum_{i=0}^{\tau-1} 1_B(\Pi_i^\alpha, \Pi_i^\beta) \right] \eta_{u(\gamma)}^\delta(dx)}{\int_{\Gamma} E_x^{\delta, u(\gamma)} [\tau] \eta_{u(\gamma)}^\delta(dx)}$$

and $\delta = \alpha, \beta$ and $B \in \sigma(\mathcal{P}(E) \times \mathcal{P}(E))$.

Proof. The verification of the continuity of K_u^3 follows from the boundedness and continuity of $F(\cdot, \cdot, \cdot)(\cdot) : \mathcal{P}(E) \times U \times \mathcal{A} \times E \setminus \partial\Gamma \rightarrow \mathbb{R}$ (cf. Theorems 1 and 6 of [18]).

If $K_u^3(\alpha, \beta, \gamma) = 0$, then by the strict positivity of p it follows that

$$E_x^{\beta, u(\gamma)} \left[\sum_{i=0}^{\tau-1} \ln \left(\frac{F(\Pi_i^\beta, u(\Pi_i^\gamma), \beta)(Y_{i+1})}{F(\Pi_i^\alpha, u(\Pi_i^\gamma), \alpha)(Y_{i+1})} \right) \right] = 0$$

for almost all $x \in \Gamma$. The map $L : \Gamma \rightarrow \mathbb{R}$, where

$$L(x) = E_x^{\beta, u(\gamma)} \left[\sum_{i=0}^{\tau-1} \ln \left(\frac{F(\Pi_i^\beta, u(\Pi_i^\gamma), \beta)(Y_{i+1})}{F(\Pi_i^\alpha, u(\Pi_i^\gamma), \alpha)(Y_{i+1})} \right) \right]$$

is continuous for each $\beta \in \mathcal{A}$, and $u \in U$ (cf. Lemma 8 of [18]), so for all $x \in \Gamma$, $L(x) = 0$. Using (A5) and Lemma 4 of [8] it follows that

$$P_x^{\beta, u(\gamma)}(\Pi_i^\alpha = \Pi_i^\beta \text{ for } i \in \{0, \dots, \tau-1\}) = 1$$

for each $x \in \Gamma$. It follows from Corollary 2 of [8] that

$$E_x^{\beta, u(\gamma)} \left[\sum_{i=0}^{\tau-1} 1_B(\Pi_i^\beta, \Pi_i^\gamma) \right] = E_x^{\alpha, u(\gamma)} \left[\sum_{i=0}^{\tau-1} 1_B(\Pi_i^\beta, \Pi_i^\gamma) \right]$$

for each B in the Borel σ -algebra of $\mathcal{P}(E) \times \mathcal{P}(E)$, $x \in \Gamma_1$, and $\eta_{u(\gamma)}^\alpha = \eta_{u(\gamma)}^\beta$. Thus $\Psi_{u(\gamma)}^\alpha = \Psi_{u(\gamma)}^\beta$. \square

4. Almost self-optimal adaptive strategies. For $\varepsilon > 0$ fixed, the controls are restricted to the finite families $\mathcal{U}^1(\varepsilon)$, $\mathcal{U}^2(\varepsilon)$, and $\tilde{\mathcal{U}}(\varepsilon)$ of ε optimal controls for Models I, II, and III, respectively. For $\varepsilon > 0$ there is a $\delta_0 > 0$ given in Proposition 3 and a δ_0 net of \mathcal{A} with a distinguished point from each element of the net $\tilde{\mathcal{A}}(\delta_0) = \{\alpha(1), \dots, \alpha(k(\delta_0))\}$.

For a randomization of an adaptive control the following subsets of \mathbb{R} are used. For $\varepsilon > 0$ let

$$(51) \quad \begin{aligned} S(\varepsilon) = \{ & \beta^i(j) : i \in \mathbb{N}, j \in \{1, 2, \dots, r(\varepsilon)\} \text{ and for each } i \in \mathbb{N} \\ & \text{there is a } j_i \in \{1, \dots, r(\varepsilon)\} \text{ such that } \beta^i(j_i) = 1 - \varepsilon/\|c\| \\ & \text{and for } j \neq j_i, \beta^i(j) = \varepsilon/[(r(\varepsilon) - 1)\|c\|] \} \end{aligned}$$

and

$$(52) \quad \begin{aligned} \tilde{S}(\varepsilon, \delta_0) = \{ & \beta^i(j, k) : i \in \mathbb{N}, j \in \{1, \dots, r(\varepsilon)\}, k \in \{1, \dots, k(\delta_0)\}, \\ & \text{and for each } i \in \mathbb{N} \text{ there are } j_i \text{ and } k_i \text{ such that } \beta^i(j_i, k_i) = 1 - \varepsilon/\|c\| \\ & \text{and for } j \neq j_i \text{ or } k \neq k_i, \beta^i(j, k) = \varepsilon/[(r(\varepsilon)k(\delta_0) - 1)\|c\|] \}, \end{aligned}$$

where $\|\cdot\|$ is the supremum norm.

The following result is a continuity property of the invariant measures for the three models and is naturally associated with Propositions 4, 5, and 6.

PROPOSITION 7. i) Consider Model I with the assumptions imposed on it. For $\varepsilon' > 0$ there is a $\delta > 0$ such that if $\{\tilde{\beta}^i(j)\} \subset S(\varepsilon)$, $\alpha, \beta \in \mathcal{A}$, $u^j \in \mathcal{U}^1(\varepsilon)$, and

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \left| \sum_{j=1}^{r(\varepsilon)} \tilde{\beta}^i(j) K_{u^j}^1(\alpha, \beta) \right| < \delta,$$

then

$$\sup_{u \in \mathcal{U}^1(\varepsilon)} \|\pi_u^\alpha - \pi_u^\beta\|_{\text{var}} < \varepsilon'.$$

ii) Consider Model II with the assumptions imposed on it. For $\varepsilon' > 0$ there is a $\delta > 0$ such that if $\{\tilde{\beta}^i(j)\} \subset S(\varepsilon)$, $\alpha, \beta \in \mathcal{A}$, $u^j \in \mathcal{U}^2(\varepsilon)$, and

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \left| \sum_{j=1}^{r(\varepsilon)} \tilde{\beta}^i(j) K_{u^j}^2(\alpha, \beta) \right| < \delta,$$

then

$$\sup_{u \in \mathcal{U}^2(\varepsilon)} \|\pi_u^\alpha - \pi_u^\beta\|_{\text{var}} < \varepsilon'.$$

iii) Consider Model III with the assumptions (A1)–(A5). For $\varepsilon' > 0$ there is a $\delta > 0$ such that if $\{\tilde{\beta}^i(j, k)\} \subset \tilde{S}(\varepsilon, \delta_0)$, $\alpha, \beta \in \mathcal{A}$, $u^j \in \tilde{\mathcal{U}}(\varepsilon)$, and

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \left| \sum_{j=1}^{r(\varepsilon)} \sum_{k=1}^{k(\delta_0)} \tilde{\beta}^i(j, k) K_{u^j}^3(\alpha, \beta, \alpha(k)) \right| < \delta,$$

then

$$\begin{aligned} & \sup_{u \in \tilde{\mathcal{U}}(\varepsilon)} \sup_{\gamma \in \tilde{\mathcal{A}}(\delta_0)} \left| \int_{\mathcal{P}(E) \times \mathcal{P}(E)} \int_E c(z, u(\nu_2)) \nu_1(dz) \Psi_{u(\gamma)}^\alpha(d\nu_1, d\nu_2) \right. \\ & \quad \left. - \int_{\mathcal{P}(E) \times \mathcal{P}(E)} \int_E c(z, u(\nu_2)) \nu_1(dz) \Psi_{u(\gamma)}^\beta(d\nu_1, d\nu_2) \right| < \varepsilon'. \end{aligned}$$

Proof. Only the verifications of i) and iii) are given because the verification of ii) is similar to that of i).

Verifying by contradiction, assume that i) is not true. Then there are sequences $(\alpha_m, m \in \mathbb{N})$ and $(\beta_m, m \in \mathbb{N})$ and $\{\tilde{\beta}^{i_m}(j)\} \subset S(\varepsilon)$ such that $\alpha_m \rightarrow \alpha$, $\beta_m \rightarrow \beta$ as $m \rightarrow \infty$,

$$(53) \quad \limsup_{m \rightarrow \infty} \liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \left| \sum_{j=1}^{r(\varepsilon)} \tilde{\beta}^{i_m}(j) K_{u^j}^1(\alpha_m, \beta_m) \right| = 0,$$

and

$$(54) \quad \sup_{u \in \mathcal{U}^1(\varepsilon)} \|\pi_u^{\alpha_m} - \pi_u^{\beta_m}\|_{\text{var}} > \varepsilon'.$$

By (53) and the definition of $S(\varepsilon)$, it follows that

$$\lim_{m \rightarrow \infty} K_{u^j}^1(\alpha_m, \beta_m) = 0$$

for each $j \in \{1, \dots, r(\varepsilon)\}$. Thus, by Proposition 4, $K_{u^j}^1(\alpha, \beta) = 0$ for $j \in \{1, \dots, r(\varepsilon)\}$ and $\pi_{u^j}^\alpha = \pi_{u^j}^\beta$ for $j \in \{1, \dots, r(\varepsilon)\}$. By (24) it follows that

$$\lim_{m \rightarrow \infty} \sup_{u \in \mathcal{U}^1(\varepsilon)} \|\pi_u^{\alpha_m} - \pi_u^\alpha\|_{\text{var}} = 0$$

and

$$\lim_{m \rightarrow \infty} \sup_{u \in \mathcal{U}^1(\varepsilon)} \|\pi_u^{\beta_m} - \pi_u^\beta\|_{\text{var}} = 0.$$

These last two equalities contradict (54). This contradiction verifies i).

Now assume that iii) is not satisfied. Then there are sequences $(\alpha_m, m \in \mathbb{N})$ and $(\beta_m, m \in \mathbb{N})$ and $\{\tilde{\beta}^{i_m}(j, k)\} \subset \tilde{S}(\varepsilon, \delta_0)$ such that $\alpha_m \rightarrow \alpha$, $\beta_m \rightarrow \beta$ as $m \rightarrow \infty$,

$$(55) \quad \limsup_{m \rightarrow \infty} \liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \left| \sum_{j=1}^{r(\varepsilon)} \sum_{k=1}^{k(\delta_0)} \tilde{\beta}^{i_m}(j, k) K_{u^j}^3(\alpha_m, \beta_m, \alpha(k)) \right| = 0$$

and

$$(56) \quad \sup_{u \in \tilde{\mathcal{U}}(\varepsilon)} \left| \int_{\mathcal{P}(E) \times \mathcal{P}(E)} \int_E c(z, u(\nu_2)) \nu_1(dz) \Psi_{u(\gamma)}^{\alpha_m}(d\nu_1, d\nu_2) - \int_{\mathcal{P}(E) \times \mathcal{P}(E)} \int_E c(z, u(\nu_2)) \nu_1(dz) \Psi_{u(\gamma)}^{\alpha}(d\nu_1, d\nu_2) \right| \geq \varepsilon'.$$

By (55) and the definition of $\tilde{S}(\varepsilon, \delta_0)$, it follows that

$$\lim_{m \rightarrow \infty} K_{u^j}^3(\alpha_m, \beta_m, \alpha(k)) = 0$$

for $j \in \{1, \dots, r(\varepsilon)\}$ and $k \in \{1, \dots, k(\delta_0)\}$. Thus, by Proposition 6, $K_{u^j}^3(\alpha, \beta, \alpha(k)) = 0$ for $j \in \{1, \dots, r(\varepsilon)\}$, $k \in \{1, \dots, k(\delta_0)\}$, and

$$\Psi_{u(\gamma)}^{\alpha} = \Psi_{u(\gamma)}^{\beta}$$

for $u \in \tilde{\mathcal{U}}(\varepsilon)$ and $\gamma \in \tilde{\mathcal{A}}(\delta_0)$. In the proof of Theorem 6 of [18] it is shown that

$$\lim_{m \rightarrow \infty} \Psi_{u(\gamma)}^{\alpha_m} = \Psi_{u(\gamma)}^{\alpha}$$

and

$$\lim_{m \rightarrow \infty} \Psi_{u(\gamma)}^{\beta_m} = \Psi_{u(\gamma)}^{\beta}$$

in the weak* topology of $\mathcal{P}(E) \times \mathcal{P}(E)$. By the continuity of c in the cost functional (11) there is a contradiction to (56). This contradiction verifies iii). \square

Fix $\varepsilon > 0$. For Models I and II let $\varepsilon' = \varepsilon/\|c\|$, and for Model III let $\varepsilon' = \varepsilon$. Using this ε' by Proposition 7, there is a $\delta > 0$ such that i), ii), and iii) are satisfied for Models I, II, and III, respectively. There is a $\bar{\delta} > 0$ such that the following are satisfied.

i) For Model I and $\alpha, \beta \in \mathcal{A}$, if $\rho_{\mathcal{A}}(\alpha, \beta) < \bar{\delta}$, then for each $u \in \mathcal{U}^1(\varepsilon)$

$$(57) \quad |K_u^1(\alpha, \beta)| < \delta/3$$

and

$$(58) \quad \|\pi_u^{\alpha} - \pi_u^{\beta}\|_{\text{var}} \leq \frac{\varepsilon}{\|c\|}.$$

ii) For Model II and $\alpha, \beta \in \mathcal{A}$, if $\rho_{\mathcal{A}}(\alpha, \beta) < \bar{\delta}$, then for each $u \in \mathcal{U}^2(\varepsilon)$

$$(59) \quad |K_u^2(\alpha, \beta)| < \delta/3$$

and

$$(60) \quad \|\pi_u^\alpha - \pi_u^\beta\|_{\text{var}} \leq \frac{\varepsilon}{\|c\|}.$$

iii) For Model III and $\alpha, \beta \in \mathcal{A}$, if $\rho_{\mathcal{A}}(\alpha, \beta) < \bar{\delta}$, then for each $u \in \tilde{\mathcal{U}}(\varepsilon)$ and $\gamma \in \tilde{\mathcal{A}}(\delta_0)$

$$(61) \quad |K_u^3(\alpha, \beta, \gamma)| < \delta/3$$

and

$$(62) \quad \left| \int_{\mathcal{P}(E) \times \mathcal{P}(E)} \int_E c(z, u(\nu_2)) \nu_1(dz) \Psi_{u(\gamma)}^{\alpha_m}(d\nu_1, d\nu_2) - \int_{\mathcal{P}(E) \times \mathcal{P}(E)} \int_E c(z, u(\nu_2)) \nu_1(dz) \Psi_{u(\gamma)}^\alpha(d\nu_1, d\nu_2) \right| \leq \varepsilon.$$

The existence of $\bar{\delta} > 0$ follows in i) from Proposition 4 and (24), in ii) from Proposition 5 and (27), (29), (30), (31), and in iii) from Proposition 6 and the continuity of the map $\Psi : \mathcal{A} \times \mathcal{A} \rightarrow \mathcal{P}(E) \times \mathcal{P}(E)$, which follows from the proof of Theorem 6 of [18].

For $\bar{\delta} > 0$ there is a finite covering of \mathcal{A} by balls of radius $\bar{\delta}$ with centers at distinguished points that is denoted $\mathcal{A}(\bar{\delta})$. For $\varepsilon > 0$ and $\delta > 0$ given above, there is a positive integer N whose existence is justified subsequently such that

i) for Model I

$$(63) \quad \sup_{x \in E} \sup_{u \in \mathcal{U}^1(\varepsilon)} \sup_{\alpha \in \mathcal{A}(\bar{\delta})} \sup_{\beta \in \mathcal{A}} \left| \frac{1}{N} \sum_{i=1}^{N-1} E_x^{\beta, u} \left[\ln \left(\frac{p(X_i, X_{i+1}, u(X_i), \beta)}{p(X_i, X_{i+1}, u(X_i), \alpha)} \right) \right] - K_u^1(\alpha, \beta) \right| < \delta/3$$

and

$$(64) \quad \sup_{x \in E} \sup_{u \in \mathcal{U}^1(\varepsilon)} \sup_{\beta \in \mathcal{A}} \left| \frac{1}{N} \sum_{i=1}^{N-1} E_x^{\beta, u} [c(X_i, u(X_i))] - \int_E c(z, u(z)) \pi_u^\beta(dz) \right| < \varepsilon;$$

ii) for Model II

$$(65) \quad \sup_{x \in \Gamma} \sup_{u \in \mathcal{U}^2(\varepsilon)} \sup_{\alpha \in \mathcal{A}(\bar{\delta})} \sup_{\beta \in \mathcal{A}} \left| E_x^{\beta, u} \left[\frac{1}{2} \int_0^{\tau_N} |\sigma^{-1}(X(s))(h(X(s), \alpha, u(X(s))) - h(X(s), \beta, u(X(s))))|^2 ds \right] (E_x^{\beta, u}[\tau_N])^{-1} - K_u^2(\alpha, \beta) \right| < \delta/3$$

and

$$(66) \quad \sup_{x \in \Gamma_1} \sup_{u \in \mathcal{U}^2(\varepsilon)} \sup_{\beta \in \mathcal{A}} \left| E_x^{\beta, u} \left[\int_0^{\tau_N} c(X(s), u(X(s))) ds \right] \cdot (E_x^{\beta, u}[\tau_N])^{-1} - \int_E c(z, u(z)) \pi_u^\beta(dz) \right| < \varepsilon;$$

iii) for Model III

$$(67) \quad \sup_{x \in \Gamma} \sup_{u \in \tilde{\mathcal{U}}(\varepsilon)} \sup_{\alpha \in \mathcal{A}(\delta)} \sup_{\gamma \in \tilde{\mathcal{A}}(\delta_0)} \sup_{\beta \in \mathcal{A}} \left| E_x^{\beta, u(\gamma)} \left[\sum_{i=0}^{\tau_N-1} \ln \left(\frac{F(\Pi_i^\beta, u(\Pi_i^\gamma), \beta)(Y_{i+1})}{F(\Pi_i^\alpha, u(\Pi_i^\gamma), \alpha)(Y_{i+1})} \right) \right] \right. \\ \left. \cdot (E_x^{\beta, u}[\tau_N])^{-1} - K_u^3(\alpha, \beta, \gamma) \right| < \delta/3$$

and

$$(68) \quad \sup_{x \in \Gamma} \sup_{u \in \tilde{\mathcal{U}}(\varepsilon)} \sup_{\gamma \in \tilde{\mathcal{A}}(\delta_0)} \sup_{\beta \in \mathcal{A}} \left| E_x^{\beta, u(\gamma)} \left[\sum_{i=0}^{\tau_N-1} \int_E c(z, u(\Pi_i^\alpha)) \Pi_i^\beta(dz) \right] (E_x^{\beta, U(\gamma)}[\tau_N])^{-1} \right. \\ \left. - \int_{\mathcal{P}(E) \times \mathcal{P}(E)} \int_E c(z, u(\nu_2)) \nu_1(dz) \Psi_{u(\gamma)}^\beta(d\nu_1, d\nu_2) \right| < \varepsilon,$$

where Ψ is given in (50).

The following three lemmas justify the existence of N in (63)–(68).

LEMMA 1. *For Model I let*

$$L = \sup_{x, y \in E} \sup_{v \in U} \sup_{\alpha, \beta \in \mathcal{A}} \left| \ln \frac{p(x, y, v, \beta)}{p(x, y, v, \alpha)} \right|.$$

Then for $N \geq 3L/(\delta(1 - \gamma_0))$ the inequality (63) is satisfied, and for $N \geq \|c\|/(\varepsilon(1 - \gamma_0))$ the inequality (64) is satisfied where γ_0 is as given in Proposition 1.

Proof. By (23) it follows that

$$\left| E_x^{\beta, u} \left[\ln \frac{p(X_i, X_{i+1}, u(X_i), \beta)}{p(X_i, X_{i+1}, u(X_i), \alpha)} \right] - K_u^1(\alpha, \beta) \right| \\ = \left| E_x^{\beta, u} \left[\int_E \ln \left(\frac{p(X_i, y, u(X_i), \beta)}{p(X_i, y, u(X_i), \alpha)} \right) p(X_i, y, u(X_i), \beta) \varphi(dy) \right] - K_u^1(\alpha, \beta) \right| \leq L\gamma_0^{i-1}$$

for $x \in E$, $\alpha, \beta \in \mathcal{A}$, $u \in \mathcal{U}(\varepsilon)$, and $i \in \mathbb{N}$. Thus for $N \geq 3L/\delta(1 - \gamma_0)$, the inequality (63) is satisfied. In a similar way by (23) it follows that the inequality (64) is satisfied for $N \geq \|c\|/\varepsilon(1 - \gamma_0)$. \square

LEMMA 2. *For Model II let*

$$L = \sup_{x \in \mathbb{R}^n} \sup_{\alpha \in \mathcal{A}} \sup_{v \in U} \|h(x, \alpha, v)^* \sigma^{-1}(x)\|,$$

$$M_1 = \sup_{x \in \Gamma_1} \sup_{\alpha \in \mathcal{A}} \sup_{u \in \mathcal{U}} E_x^{\alpha, u}[\tau],$$

and

$$M_2 = \inf_{x \in \Gamma_1} \inf_{\alpha \in \mathcal{A}} \inf_{u \in \mathcal{U}} E_x^{\alpha, u}[\tau].$$

For

$$N \geq \frac{12}{\delta} L^2 M_1 \frac{1}{M_2(1 - \gamma_0)} \left(1 + \frac{M_1}{M_2} \right),$$

the inequality (65) is satisfied, and for

$$N \geq \frac{1}{\varepsilon} \|c\| M_1 \frac{1}{M_2(1-\gamma_0)} \left(1 + \frac{M_1}{M_2}\right),$$

where γ_0 is given in Proposition 2, the inequality (66) is satisfied.

Proof. By (28) it follows that

$$\begin{aligned} & \left| \frac{1}{i} E_x^{\beta,u} \left[\int_0^{\tau_{i+1}} |\sigma^{-1}(X(s))(h(X(s), \alpha, u(X(s))) - h(X(s), \beta, u(X(s))))|^2 ds \right] \right. \\ & \left. - \int_{\Gamma_1} E_x^{\beta,u} \left[\int_0^\tau |\sigma^{-1}(X(s))(h(X(s), \alpha, u(X(s))) - h(X(s), \beta, u(X(s))))|^2 ds \right] \eta_u^\beta(dz) \right| \\ & \leq \frac{1}{i} \sum_{j=0}^i \gamma_0^j 4L^2 M_1 = \frac{1}{i} \frac{1 - \gamma_0^{i+1}}{1 - \gamma_0} 4L^2 M_1 \end{aligned}$$

and

$$\left| \frac{1}{i} E_x^{\beta,u}[\tau_{i+1}] - \int_{\Gamma_1} E_z^{\beta,u}[\tau] \eta_u^\beta(dz) \right| \leq \frac{1}{i} \frac{1 - \gamma_0^{i+1}}{1 - \gamma_0} M_1$$

for $x \in \Gamma_1$, $\beta \in \mathcal{A}$, and $u \in \mathcal{U}$.

Combining the above two inequalities, (65) is satisfied for N stated in the lemma.

In a similar way, (66) is verified. \square

LEMMA 3. For Model III let

$$L_1 = \sup_{x,y \in E} \sup_{v \in U} \sup_{\alpha \in \mathcal{A}} p(x, y, v, \alpha),$$

$$L_2 = \inf_{x,y \in E} \inf_{v \in U} \inf_{\alpha \in \mathcal{A}} p(x, y, v, \alpha),$$

$$M_1 = \sup_{x \in \Gamma} \sup_{u \in \mathcal{U}(\varepsilon)} \sup_{\gamma \in \mathcal{A}(\delta_0)} \sup_{\beta \in \mathcal{A}} E_x^{\beta,u(\gamma)}[\tau],$$

and

$$M_2 = \inf_{x \in \Gamma} \inf_{u \in \mathcal{U}(\varepsilon)} \inf_{\gamma \in \mathcal{A}(\delta_0)} \inf_{\beta \in \mathcal{A}} E_x^{\beta,u(\gamma)}[\tau].$$

For

$$N \geq \frac{3}{\delta} \ln \left(\frac{L_1}{L_2} \right) M_1 \frac{1}{M_2(1-\gamma_0)} \left(1 + \frac{M_1}{M_2}\right),$$

the inequality (67) is satisfied, and for

$$N \geq \frac{\|c\|}{\varepsilon} M_1 \frac{1}{(1-\gamma_0)M_2} \left(1 + \frac{M_1}{M_2}\right),$$

the inequality (68) is satisfied where γ_0 is given in Proposition 3.

Proof. As in the verification of Lemma 2 it follows by (40) that

$$\left| \frac{1}{i} E_x^{\beta, u(\gamma)} \left[\sum_{j=0}^{\tau_{i+1}-1} \ln \frac{F(\Pi_j^\beta, u(\Pi_j^\gamma), \beta)(Y_{j+1})}{F(\Pi_j^\alpha, u(\Pi_j^\gamma), \alpha)(Y_{j+1})} \right] \right. \\ \left. - \int_\Gamma E_x^{\beta, u(\gamma)} \left[\sum_{j=0}^{\tau-1} \ln \frac{F(\Pi_j^\beta, u(\Pi_j^\gamma), \beta)(Y_{j+1})}{F(\Pi_j^\alpha, u(\Pi_j^\gamma), \alpha)(Y_{j+1})} \right] \cdot \eta_{u(\gamma)}^\beta(dx) \right| \leq \frac{1}{i} \frac{1 - \gamma_0^{i+1}}{1 - \gamma_0} \ln \left(\frac{L_1}{L_2} \right) M_1$$

and

$$\left| \frac{1}{i} E_x^{\beta, u(\gamma)}[\tau_{i+1}] - \int_\Gamma E_z^{\beta, u(\gamma)}[\tau] \eta_{u(\gamma)}^\beta(dz) \right| \leq \frac{1}{i} \frac{1 - \gamma_0^{i+1}}{1 - \gamma_0} M_1$$

for $x \in \Gamma$, $\beta \in \mathcal{A}$, $\gamma \in \tilde{\mathcal{A}}(\delta_0)$, $u \in \tilde{\mathcal{U}}(\varepsilon)$. These inequalities imply the inequalities for N for which (67) and (68) are satisfied. \square

Now the construction of the almost self-optimal controls can be completed. Again, it is subdivided into the three models.

i) For Model I let $\hat{\alpha}_{jN}$ be a maximizer of

$$(69) \quad L_{jN}^1(\alpha) = \sum_{i=0}^{jN-1} \ln p(X_i, X_{i+1}, v_i, \alpha)$$

over $\alpha \in \mathcal{A}(\bar{\delta})$, where v_i is the control at time i . The control v_i is a randomized certainty equivalence control. For $i \in \{jN : j \in \mathbb{N}\}$, choose the control $u_{jN} \in \mathcal{U}^1(\varepsilon)$ randomly among $(u^k, k = 1, \dots, r(\varepsilon))$ as

$$(70) \quad P(u_{jN} = u^{k_0} | X(0), \dots, X(jN)) = 1 - \frac{\varepsilon}{\|c\|},$$

where u^{k_0} is the almost optimal control for $\alpha = \hat{\alpha}_{jN}$ in $\mathcal{U}^1(\varepsilon)$ and

$$(71) \quad P(u_{jN} = u^k | X(0), \dots, X(jN)) = \frac{\varepsilon}{(r(\varepsilon) - 1)\|c\|}$$

for $k = \{1, \dots, r(\varepsilon)\} \setminus \{k_0\}$. The control u_{jN} is also used at the times $jN+1, \dots, (j+1)N-1$; that is,

$$(72) \quad v_i = u_{jN}(X_i)$$

for $i = \{jN, \dots, (j+1)N-1\}$.

ii) For Model II, let $\hat{\alpha}(\tau_{jN})$ be a minimizer of

$$(73) \quad L^2(\tau_{jN}) = \int_0^{\tau_{jN}} |\sigma^{-1}(X(s))(h(X(s), \alpha, v(s)) - h(X(s), \alpha^0, v(s)))|^2 ds$$

over $\alpha \in \mathcal{A}(\delta)$. The control in $[\tau_{jN}, \tau_{(j+1)N})$ is $u(\tau_{jN}) \in \mathcal{U}^2(\varepsilon)$, that is, a randomized certainty equivalence control such that

$$(74) \quad P(u(\tau_{jN}) = u^{k_0} | X(s), 0 \leq s \leq \tau_{jN}) = 1 - \frac{\varepsilon}{\|c\|}$$

and

$$(75) \quad P(u(\tau_{jN}) = u^k | X(s), 0 \leq s \leq \tau_{jN}) = \frac{\varepsilon}{(r(\varepsilon) - 1)\|c\|}$$

for $k = \{1, \dots, r(\varepsilon)\} \setminus \{k_0\}$ and u^{k_0} is almost optimal for $\hat{\alpha}(\tau_{jN})$.

iii) For Model III let $\hat{\alpha}(\tau_{jN})$ be a maximizer of

$$(76) \quad L_{\tau_{jN}}^3(\alpha) = \sum_{i=0}^{\tau_{jN}-1} \ln F(\Pi_i^\alpha, v_i, \alpha)(Y_{i+1})$$

over $\alpha \in \mathcal{A}(\bar{\delta})$. The controls $\hat{v}(\tau_{jN}), \hat{v}(\tau_{jN+1}), \dots, \hat{v}(\tau_{(j+1)N-1})$ are selected by a randomized certainty equivalence rule such that

$$(77) \quad P(\hat{v}(\tau_{jN}) = u^{k_0}(\Pi_{\tau_{jN}}^{\alpha(l_0)}), \dots, \hat{v}(\tau_{(j+1)N-1}) = u^{k_0}(\Pi_{\tau_{(j+1)N-1}}^{\alpha(l_0)}) | \mathcal{Y}(\tau_{jN})) = 1 - \frac{\varepsilon}{\|c\|}$$

and

$$(78) \quad \begin{aligned} P(\hat{v}(\tau_{jN}) = u^k(\Pi_{\tau_{jN}}^{\alpha(l)}), \dots, \hat{v}(\tau_{(j+1)N-1}) = u^k(\Pi_{\tau_{(j+1)N-1}}^{\alpha(l)}) | \mathcal{Y}(\tau_{jN})) \\ = \frac{\varepsilon}{(r(\varepsilon)k(\varepsilon) - 1)\|c\|}, \end{aligned}$$

where $k \in \{1, \dots, r(\varepsilon)\} \setminus \{k_0\}$, $j \in \{1, \dots, \text{card}(\tilde{\mathcal{A}}(\delta_{k_0}))\}$, and $u^{k_0}(\Pi_{\tau_{jN}}^{\alpha(l_0)})$ is almost optimal for $\hat{\alpha}(\tau_{jN})$.

Let $(\hat{v}_i, i \in \mathbb{N})$ or $(\hat{v}(s), s \geq 0)$ be the discrete or the continuous time randomized certainty equivalence control defined in i), ii), or iii) above. Let $(\bar{\beta}_i, i \in \mathbb{N})$, $(\bar{\beta}(s), s \geq 0)$, and $(\tilde{\beta}_i, i \in \mathbb{N})$ be processes with values in $\{1, \dots, r(\varepsilon)\}$ for the first two processes and in $\{(j, k) : j = 1, \dots, r(\varepsilon) \text{ and } k = 1, \dots, k(\delta_0)\}$ for the third process such that the first two processes correspond to the index of the control in $\mathcal{U}^i(\varepsilon)$, $i = 1, 2$, at each time and the third process (for Model III) is the index of the control function and the index of the element of $\tilde{\mathcal{A}}(\delta_0)$.

The following result is the almost self-optimality of the randomized certainty equivalence control for the three Models I, II, and III.

THEOREM 1. *Let $\varepsilon > 0$ be fixed. Let I^i be the pathwise cost functional for $i = 1, 2, 3$ given by (2), (8), and (11), respectively, for Models I, II, and III, respectively, and let $\lambda_i(\alpha^0)$ be the optimal cost for $i = 1, 2, 3$ for parameter α^0 . Let $(\hat{v}_i, i \in \mathbb{N})$ and $(\hat{v}(s), s \geq 0)$ be the randomized certainty equivalence controls given above. For the Models I, II, and III the following inequalities are satisfied:*

$$(79) \quad \text{(i) } I^1((\hat{v}_n, n \in \mathbb{N})) \leq \lambda_1(\alpha^0) + 6\varepsilon \quad a.s.,$$

$$(80) \quad \text{(ii) } I^2((\hat{v}(s), s \geq 0)) \leq \lambda_2(\alpha^0) + 6\varepsilon \quad a.s.,$$

$$(81) \quad \text{(iii) } I^3((\hat{v}_n, n \in \mathbb{N})) \leq \lambda_3(\alpha^0) + 6\varepsilon \quad a.s.$$

Proof. Initially consider Model I. By the definition of $\hat{\alpha}_{jN}$, it follows that

$$(82) \quad \sum_{i=0}^{jN-1} \ln \frac{p(X_i, X_{i+1}, \hat{v}_i, \hat{\alpha}_{jN})}{p(X_i, X_{i+1}, \hat{v}_i, \alpha^0)} \geq \sum_{i=0}^{jN-1} \ln \frac{p(X_i, X_{i+1}, \hat{v}_i, \bar{\alpha}^0)}{p(X_i, X_{i+1}, \hat{v}_i, \alpha^0)},$$

where $\bar{\alpha}^0$ is an element in $\mathcal{A}(\bar{\delta})$ of minimum distance to α^0 .

By the law of large numbers for martingales for $\alpha \in \mathcal{A}(\bar{\delta})$,

$$(83) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \left(\sum_{i=0}^{nN-1} \ln \frac{p(X_i, X_{i+1}, \hat{v}_i, \alpha)}{p(X_i, X_{i+1}, \hat{v}_i, \alpha^0)} \right. \\ \left. - \sum_{i=0}^{n-1} E^{\alpha^0} \left[\sum_{j=iN}^{(i+1)N-1} \ln \frac{p(X_i, X_{i+1}, \hat{v}_i, \alpha)}{p(X_i, X_{i+1}, \hat{v}_i, \alpha^0)} \middle| X_0, \dots, X_{iN}, \bar{\beta}_0, \dots, \bar{\beta}_{iN} \right] \right) = 0 \quad \text{a.s.},$$

$$(84) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \left(\sum_{i=0}^{nN-1} c(X_i, \hat{v}_i) \right. \\ \left. - \sum_{i=0}^{n-1} E^{\alpha^0} \left[\sum_{j=iN}^{(i+1)N-1} c(X_j, \hat{v}_j) \middle| X_0, \dots, X_{iN}, \bar{\beta}_0, \dots, \bar{\beta}_{iN} \right] \right) = 0 \quad \text{a.s.},$$

$$(85) \quad \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \left(K_{u_{iN}}^1(\alpha, \alpha^0) - \sum_{j=1}^{r(\varepsilon)} \beta^i(j) K_{u^j}^1(\alpha, \alpha^0) \right) = 0 \quad \text{a.s.},$$

and

$$(86) \quad \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} J^{\hat{\alpha}_{iN}, 1}((u_{iN}(X_l), l \in \mathbb{N})) \\ - \sum_{j=1}^{r(\varepsilon)} \beta^i(j) J^{\hat{\alpha}_{iN}, 1}((u^j(X_l), l \in \mathbb{N})) = 0 \quad \text{a.s.},$$

where $\beta^i(j) = P(\bar{\beta}_i = j | X_0, \dots, X_{iN})$, and in the last equality the control functions u_{iN} and u^j are used and their costs are evaluated. Let \mathcal{N} be a null set such that the above four equalities are satisfied on \mathcal{N}^c . Let $F(\bar{\delta}) = \{\alpha \in \mathcal{A}(\bar{\delta}) : \text{there is an } \omega \in \Omega \setminus \mathcal{N} \text{ such that } \alpha \text{ is a frequent point of } (\hat{\alpha}_{jN}(\omega))\}$. In many subsequent expressions, the random variables are evaluated at some $\omega \in \Omega \setminus \mathcal{N}$ but this evaluation is suppressed for notational convenience. If $\alpha \in F(\bar{\delta})$ then for a corresponding $\omega \in \Omega \setminus \mathcal{N}$ it follows from (82), (83) that

$$(87) \quad \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} E_{X_{iN}}^{\alpha^0, u_{iN}} \left[\sum_{j=iN}^{(i+1)N-1} \ln \frac{p(X_j, X_{j+1}, u_{iN}(X_j), \alpha^0)}{p(X_j, X_{j+1}, u_{iN}(X_j), \alpha)} \right] \\ \geq \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} E_{X_{iN}}^{\alpha^0, u_{iN}} \left[\sum_{j=iN}^{(i+1)N-1} \ln \frac{p(X_j, X_{j+1}, u_{iN}(X_j), \alpha^0)}{p(X_j, X_{j+1}, u_{iN}(X_j), \bar{\alpha})} \right]$$

for each $\bar{\alpha} \in \mathcal{A}(\bar{\delta})$. By (63) it follows that

$$(88) \quad \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} K_{u_{iN}}^1(\alpha, \alpha^0) + \frac{\delta}{3} \geq \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} K_{u_{iN}}^1(\bar{\alpha}, \alpha^0) - \frac{\delta}{3}.$$

Thus, by (57),

$$(89) \quad \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} K_{u_{iN}}^1(\alpha, \alpha^0) \geq -\delta,$$

and by (85),

$$(90) \quad \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \sum_{j=1}^{r(\varepsilon)} \beta^i(j) K_{u_j}^1(\alpha, \alpha^0) \geq -\delta.$$

Therefore, by Proposition 7 with $\varepsilon^1 = \varepsilon/\|c\|$,

$$(91) \quad \sup_{u \in \mathcal{U}^1(\varepsilon)} \|\pi_u^\alpha - \pi_u^{\alpha^0}\|_{\text{var}} < \frac{\varepsilon}{\|c\|}.$$

By (84)

$$(92) \quad I^1((\hat{v}_n, n \in \mathbb{N})) = \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} E_{X_{iN}}^{\alpha^0, u_{iN}} \left[\sum_{j=iN}^{(i+1)N-1} c(X_j, u_{iN}(X_j)) \right] \quad \text{a.s.},$$

so by (64),

$$(93) \quad I^1((\hat{v}_n, n \in \mathbb{N})) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \int_E c(z, u_{iN}(z)) \pi_{u_{iN}}^{\alpha^0}(\mathrm{d}z) + \varepsilon \quad \text{a.s.}$$

For $\omega \in \Omega \setminus \mathcal{N}$ it follows from (93) that

$$(94) \quad I^1((\hat{v}_n, n \in \mathbb{N})) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} 1_{F(\bar{\delta})}(\hat{\alpha}_{iN}) \int_E c(z, u_{iN}(z)) \pi_{u_{iN}}^{\alpha^0}(\mathrm{d}z) + \varepsilon.$$

For $\alpha \in F(\bar{\delta})$, (91) is satisfied, so for $\omega \in \Omega \setminus \mathcal{N}$

$$(95) \quad \begin{aligned} I^1((\hat{v}_n, n \in \mathbb{N})) &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} 1_{F(\bar{\delta})}(\hat{\alpha}_{iN}) \int_E c(z, u_{iN}(z)) \pi_{u_{iN}}^{\hat{\alpha}_{iN}}(\mathrm{d}z) + 2\varepsilon \\ &= \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} J^{\hat{\alpha}_{iN}, 1}((u_{iN}(X_l), l \in \mathbb{N})) + 2\varepsilon. \end{aligned}$$

For $\omega \in \Omega \setminus \mathcal{N}$ it follows by (86) that

$$(96) \quad I^1((\hat{v}_n, n \in \mathbb{N})) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \sum_{j=1}^{r(\varepsilon)} \beta_i(j) J^{\hat{\alpha}_{iN}, 1}(u^j(X_l), l \in \mathbb{N}) + 2\varepsilon.$$

Using the definition of $\beta^i(j)$ and (25) it follows that for $\omega \in \Omega \setminus \mathcal{N}$

$$\begin{aligned} I^1((\hat{v}_n, n \in \mathbb{N})) &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \left(1 - \frac{\varepsilon}{\|c\|}\right) \lambda_1(\hat{\alpha}_{iN}) + 3\varepsilon \\ (97) \quad &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \lambda_1(\hat{\alpha}_{iN}) + 3\varepsilon \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} 1_{F(\bar{\delta})}(\hat{\alpha}_{iN}) \lambda_1(\hat{\alpha}_{iN}) + 3\varepsilon. \end{aligned}$$

By (25) and (91) it follows that for $\alpha \in F(\bar{\delta})$

$$(98) \quad |\lambda_1(\alpha) - \lambda_1(\alpha^0)| \leq 2\varepsilon + \sup_{u \in \mathcal{U}^1(\varepsilon)} \left| \int_E c(x, u(x)) (\pi_u^\alpha(dx) - \pi_u^{\alpha^0}(dx)) \right| \leq 3\varepsilon.$$

Thus for $\omega \in \Omega \setminus \mathcal{N}$

$$I^1((\hat{v}_n, n \in \mathbb{N})) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} 1_{F(\bar{\delta})}(\hat{\alpha}_{iN}) \lambda_1(\alpha^0) + 6\varepsilon = \lambda_1(\alpha^0) + 6\varepsilon.$$

The inequality (79) is verified.

Now consider Model III. By the construction of $\hat{\alpha}(\tau_{jN})$ it follows that

$$(99) \quad \sum_{i=0}^{\tau_{jN}-1} \ln \frac{F(\Pi_i^{\hat{\alpha}_{\tau_{jN}}}, \hat{v}_i, \hat{\alpha}_{\tau_{jN}})(Y_{i+1})}{F(\Pi_i^{\alpha^0}, \hat{v}_i, \alpha^0)(Y_{i+1})} \geq \sum_{i=0}^{\tau_{jN}-1} \ln \frac{F(\Pi_i^{\bar{\alpha}_i}, \hat{v}_i, \bar{\alpha}^0)(Y_{i+1})}{F(\Pi_i^{\alpha^0}, \hat{v}_i, \alpha^0)(Y_{i+1})}.$$

In analogy to (83)–(86), by the law of large numbers for martingales it follows that for $\alpha \in \mathcal{A}(\bar{\delta})$,

$$\begin{aligned} &\lim_{n \rightarrow \infty} \frac{1}{n} \left(\sum_{i=0}^{\tau_{nN}-1} \ln \frac{F(\Pi_i^{\bar{\alpha}_i}, \hat{v}_i, \alpha^0)(Y_{i+1})}{F(\Pi_i^{\alpha^0}, \hat{v}_i, \alpha^0)(Y_{i+1})} \right. \\ (100) \quad &\left. - \sum_{i=0}^{\tau_{nN}-1} E^{\alpha^0} \left[\sum_{j=\tau_{iN}}^{\tau_{(i+1)N}-1} \ln \frac{F(\Pi_j^{\bar{\alpha}_j}, \hat{v}_j, \alpha^0)(Y_{j+1})}{F(\Pi_j^{\alpha^0}, \hat{v}_j, \alpha^0)(Y_{j+1})} \middle| \mathcal{Y}(\tau_{iN}), \bar{\beta}_0, \dots, \bar{\beta}_{\tau_{iN}} \right] \right) = 0 \quad \text{a.s.}, \end{aligned}$$

$$(101) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \left(\tau_{nN} - \sum_{i=0}^{n-1} E^{\alpha^0} [\tau_{(i+1)N} - \tau_{iN} | \mathcal{Y}(\tau_{iN}), \bar{\beta}_0, \dots, \bar{\beta}_{\tau_{iN}}] \right) = 0 \quad \text{a.s.},$$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \left(\sum_{i=0}^{\tau_{nN}} c(X_i, \hat{v}_i) \right)$$

$$(102) \quad - \sum_{i=0}^{n-1} E^{\alpha^0} \left[\sum_{j=\tau_{iN}}^{\tau_{(i+1)N}-1} \int_E c(z, \hat{v}_i) \Pi_i^{\alpha^0}(dz) | \mathcal{Y}(\tau_{iN}), \bar{\beta}_0, \dots, \bar{\beta}_{\tau_{iN}} \right] = 0 \quad \text{a.s.},$$

(103)

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \left(K_{u_{\tau_{iN}}}^3(\alpha, \alpha^0, \tilde{\alpha}_{\tau_{iN}}) - \sum_{j=1}^{r(\varepsilon)} \sum_{k=1}^{k(\delta_0)} \beta^i(j, k) K_{u^j}^3(\alpha, \alpha^0, \alpha(k)) \right) = 0 \quad \text{a.s.},$$

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} (J^{\hat{\alpha}_{\tau_{iN}}, 3}(u_{\tau_{iN}}(\Pi_l^{\tilde{\alpha}_{\tau_{iN}}}), l \in \mathbb{N}))$$

$$(104) \quad - \sum_{j=1}^{r(\varepsilon)} \sum_{k=1}^{k(\delta_0)} \beta^i(j, k) (J^{\hat{\alpha}_{\tau_{iN}}, 3}(u^j(\Pi_l^{\tilde{\alpha}_{\tau_{iN}}}), l \in \mathbb{N})) = 0 \quad \text{a.s.},$$

where $\tilde{\alpha}_{\tau_{iN}}$ is the element of $\tilde{\mathcal{A}}(\delta_0)$ chosen at time τ_{iN} in the construction of the control $\hat{v}_{\tau_{iN}}$,

$$\beta_i(j, k) = P(\tilde{\beta}_{\tau_{iN}} = (j, k) | \mathcal{Y}(\tau_{iN})),$$

and $J^{\alpha, 3}$ is the evaluation of the average cost for $\alpha \in \mathcal{A}$. If $\alpha \in \mathcal{A}(\bar{\delta})$ for some $\omega \in \Omega \setminus \mathcal{N}$ is a frequent point of the estimation, then similar to (87)–(90), by (62), (67), (99) it follows that

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \sum_{j=1}^{r(\varepsilon)} \sum_{k=1}^{k(\delta_0)} \beta^i(j, k) K_{u^j}^3(\alpha, \alpha^0, \alpha(k)) \geq -\delta,$$

so by Proposition 7

$$(105) \quad \sup_{u \in \tilde{\mathcal{U}}(\varepsilon)} \sup_{\beta \in \tilde{\mathcal{A}}(\delta_0)} \left| \int_{\mathcal{P}(E) \times \mathcal{P}(E)} \int_E c(z, u(\nu_2)) \nu_1(dz) \Psi_{u(\beta)}^\alpha(d\nu_1, d\nu_2) \right. \\ \left. - \int_{\mathcal{P}(E) \times \mathcal{P}(E)} \int_E c(z, u(\nu_2)) \nu_1(dz) \Psi_{u(\beta)}^{\alpha^0}(d\nu_1, d\nu_2) \right| < \varepsilon.$$

By (68), (101), (102) it follows that

$$J^{\alpha^0, 3}((\hat{v}_n, n \in \mathbb{N})) \\ \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \int_{\mathcal{P}(E) \times \mathcal{P}(E)} \int_E c(z, u(\nu_2)) \nu_1(dz) \Psi_{u_{\tau_{iN}}(\tilde{\alpha}_{\tau_{iN}})}^{\alpha^0}(d\nu_1, d\nu_2) + \varepsilon.$$

Let $F(\bar{\delta})$ be the set given by $F(\bar{\delta}) = \{\alpha \in \mathcal{A}(\bar{\delta}) : \text{there is an } \omega \in \Omega \setminus \mathcal{N} \text{ such that } \alpha \text{ is a frequent point of } \hat{\alpha}_{\tau_{iN}}(\omega)\}$. By (104), (105) it follows that

$$I^3((\hat{v}_n, n \in \mathbb{N})) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} 1_{F(\delta)}(\hat{\alpha}_{\tau_{iN}})$$

$$\cdot \int_{\mathcal{P}(E) \times \mathcal{P}(E)} \int_E c(z, u(\nu_2)) \nu_1(dz) \Psi_{u_{\tau_{iN}}(\tilde{\alpha}_{\tau_{iN}})}^{\hat{\alpha}_{\tau_{iN}}}(d\nu_1, d\nu_2) + 2\varepsilon,$$

so by (104)

$$I^3((\hat{v}_n, n \in \mathbb{N})) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \sum_{j=1}^{r(\varepsilon)} \sum_{k=1}^{k(\delta_0)} \beta^i(j, k) \\ \cdot J^{\hat{\alpha}_{\tau_{iN}}, 3}(u^j(\pi_l^{\alpha(k)}), l \in \mathbb{N}) + 2\varepsilon.$$

Similar to (97), it follows by (104) that

$$I^3((\hat{v}_n, n \in \mathbb{N})) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} 1_{F(\bar{\delta})}(\hat{\alpha}_{\tau_{iN}}) \lambda_3(\hat{\alpha}_{\tau_{iN}}) + 3\varepsilon.$$

For $\alpha \in F(\bar{\delta})$ it follows from (41), (104) that

$$|\lambda_3(\alpha) - \lambda_3(\alpha^0)| \leq 3\varepsilon$$

and

$$I^3((\hat{v}_n, n \in \mathbb{N})) \leq \lambda_3(\alpha^0) + 6\varepsilon \quad \text{a.s.}$$

This verifies the inequality (81).

The verification of the inequality (80) for Model II is similar to the verification of (79) and (81), and is thereby omitted. \square

5. Some other adaptive algorithms. The existence of a finite family of almost optimal controls that is shown in section 2 can be used in the construction of some other algorithms. Three such algorithms are i) maximum likelihood estimation with forcing, ii) cost watching with forcing, and iii) cost watching with randomization. The forcing algorithms (cf., e.g., [2]) are based on the forced use of all of the almost optimal controls successively at times $(T_n, n \in \mathbb{N})$ such that

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} 1_{\{T_j, j \in \mathbb{N}\}}(i) = 0,$$

so that the forcing does not affect the value of the cost functional. The notion of cost watching is to compare the average costs incurred for each of the almost optimal controls. For cost watching during the nonforcing times, a control that has minimal average cost is used. For cost watching with randomization the control for which the current average cost is minimal is chosen with probability almost 1, and the other controls are chosen with small probability. It seems that the algorithms with forcing should converge slowly, but they have a simple construction. The algorithm given in section 4 is more complicated and requires some continuity properties of the invariant measures with respect to the information. A comparative analysis of the above algorithms requires further study.

REFERENCES

- [1] R. AGRAWAL, *Minimizing the learning loss in adaptive control of Markov chains under the weak accessibility condition*, J. Appl. Prob., 28 (1991), pp. 779–790.
- [2] R. AGRAWAL AND D. TENEKETZIS, *Certainty equivalence control with forcing: Revisited*, Systems Control Lett., 13 (1989), pp. 405–412.

- [3] A. BENSOUSSAN, *Perturbation Methods in Optimal Control*, J. Wiley, New York, 1988.
- [4] V. S. BORKAR, *Self-tuning control of diffusions without the identifiability condition*, J. Optim. Theory Appl., 68 (1991), pp. 117–138.
- [5] J. L. DOOB, *Stochastic Processes*, J. Wiley, New York, 1953.
- [6] T. E. DUNCAN, B. PASIK-DUNCAN, AND L. STETTNER, *Almost self-optimizing strategies for the adaptive control of diffusion processes*, J. Optim. Theory Appl., 81 (1994), pp. 479–507.
- [7] T. E. DUNCAN, B. PASIK-DUNCAN, AND L. STETTNER, *On the ergodic and the adaptive control of stochastic delay systems*, J. Optim. Theory Appl., 81 (1994), pp. 509–531.
- [8] T. E. DUNCAN, B. PASIK-DUNCAN, AND L. STETTNER, *Adaptive control of a partially observed discrete time Markov process*, J. Appl. Math. Optim., 1998, to appear.
- [9] E. FERNANDEZ-GAUCHERAND, A. ARAPOSTATHIS, AND S. I. MARCUS, *Analysis of an adaptive control scheme for a partially observed Markov chain*, IEEE Trans. Automat. Control, 38 (1993), pp. 987–993.
- [10] O. HERNANDEZ-LERMA, *Adaptive Markov Control Processes*, Springer-Verlag, Berlin, 1989.
- [11] S. KULLBACK AND R. A. LEIBLER, *On information and sufficiency*, Ann. Math. Stat., 22 (1951), pp. 79–86.
- [12] S. KULLBACK, *Information Theory and Statistics*, J. Wiley, New York, 1959.
- [13] P. R. KUMAR AND A. BECKER, *A new family of adaptive controllers for Markov chains*, IEEE Trans. Automat. Control, 27 (1982), pp. 137–146.
- [14] H. J. KUSHNER, *Approximation methods for minimum average cost per unit time problems with a diffusion model*, in Approximate Solutions of Random Equations, A. T. Bharucha-Reid, ed., North-Holland, Amsterdam, 1979, pp. 109–126.
- [15] W. J. Runggaldier and L. Stettner, *Approximations of Discrete Time Partially Observed Control Problems*, Appl. Math. Monographs 6, C.N.R., Pisa, 1994.
- [16] L. J. SAVAGE, *The Foundations of Statistics*, J. Wiley, New York, 1954.
- [17] L. STETTNER, *On nearly self-optimizing strategies for a discrete-time uniformly ergodic adaptive model*, J. Appl. Math. Optim., 27 (1993), pp. 161–177.
- [18] L. STETTNER, *Ergodic control of Markov process with mixed observation structure*, Dissertationes Math., 341 (1995), pp. 1–36.
- [19] D. W. STROOCK AND S. R. S. VARADHAN, *Multidimensional Diffusion Processes*, Springer-Verlag, New York, 1979.